

Copyright (c) 1993 - 2000 Compugen Ltd. GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 5, 2002, 13:47:39 ; Search time 3997.53 Seconds (without alignments) 1376.994 Million cell updates/sec

Title: US-08-153-397a-1

Perfect score: 3962

Sequence: 1 CGGGCTGAGACTGGGTTA.....AAAAAAACCGGAATT 3962

Scoring table: IDENTITY\_NUC ; Gapopen 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 0

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: em\_lestba:\*

2: em\_esthun:\*

3: em\_estin:\*

4: em\_lestnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_lestro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_gss:\*

13: em\_gss\_hum:\*

14: em\_gss\_inv:\*

15: em\_gss\_pln:\*

16: em\_gss\_virt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

RESULT	1	ALIGNMENTS
LOCUS	BC006836	2633 bp mRNA linear HTC 12-JUL-2001
DEFINITION	Mus musculus, Similar to disocidin domain receptor family, member	
ACCESSION	BC006836	
VERSION	BC006836.1	
KEYWORDS	HTC	
SOURCE	house mouse	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	Mammalia; Butheira; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE	1 (bases 1 to 2633)	
JOURNAL	Submitted (27-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	
COMMENT	Contact: MGC help desk Email: cgbpbs+e@mail.nih.gov	
Tissue Procurement	Jeffrey Green M.D.	
CDNA Library Preparation	Life Technologies, Inc.	
CDNA Library Arrayed by	The I.M.A.G.E. Consortium (L1NL)	
DNA Sequencing by	Baylor College of Medicine Human Genome Sequencing Center	
Center code	BCH-HGSC	
Web site	<a href="http://www.hgsc.bcm.tmc.edu/cdna/">http://www.hgsc.bcm.tmc.edu/cdna/</a>	
Contact	<a href="mailto:villalon@bcm.tmc.edu">villalon@bcm.tmc.edu</a>	
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.,		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/L1NL at: <http://image.lnl.gov>

Series: IRAK Plate: 16 Row: 0 Column: 14

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis



JOURNAL	
COMMENT	Genoscope
CONTACT	Centre National de sequencage
GENOSCOPE	BP 191 91006 EVRY cedex - France
EMAIL	segref@genoscope.cns.fr. Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> .
FEATURES	Location/Qualifiers
source	1. 1.1010 /organism="Homo sapiens" /db_xref="Taxon:9606" /clone="CSDP033Y015" /clone_id="LTL.FL013.FBnl" /dev_stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)" /lab_host="DHLB" /note="Organ: fetal brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dN) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA. Fax: (1) 301 610 8371 Email: <a href="http://fulllength.invitrogen.com">http://fulllength.invitrogen.com</a> "
BASE COUNT	191 a 321 c 298 g 199 t 1 others
ORIGIN	
Query	Match 24.2% Score 960; DB 9; Length 1010; Best Local Similarity 99.5%; PPrd. No. 1.2e-163; Matches 983; Conservative 1; Mismatches 1; Indels 3; Gaps 2;
Db	2220 CGCTGAACTTGTGCAATCGGACCTCGCACCGGAACATCTTGGTTGGGAAATTCA 2279
Qy	2674 CCATCAAATTCGCAACTTGGCATGAGCGGAACTCTATGCTGGGGACTTACCGTG 2733
Db	2280 CCATCAAATTCGCAACTTGGCATGAGCGGAACTCTATGCTGGGGACTTACCGTG 2339
Qy	2734 TCGAGGGCGGCACTGCTGCCATCGTAGTGGAGCTGGAGTCATCTCATGGGA 2793
Db	2340 TCCAGGGCGGGCTGCCATCGTAGTGGAGCTGGAGTCATCTCATGGGA 2399
Qy	2794 AGTCACGCTGGCACTGGCTGGGGCTGGGGCTGGGGAGTGTGATG 2853
Db	2400 AGTCACACAGCCAGTGAAGTGGCCCTGGGAGTGTGATG 2459
Qy	2854 TCTTAGGCCAGCCTTGGGAGCTACCGACAGGGTCATCGAGAACGGGGG 2913
Db	2460 TCTCAGGCCAGCCTTGGGAGCTACAGATGAGGTATCGAGAACGGGG 2519
Qy	2914 AGTCCTGGACAGGCCGGAGCTGGGGAGTGTGACCTGGGGCTGGGGAGTGTG 2973
Db	2520 AGTCTCAGGCCAGGCCGGAGTGTGACCTGGGGCTGGGGAGTGTG 2579
Qy	2974 GCCTATGAGCTGATCTGGGGCTGAGCCGGAGCTGACGAGCCAC 3027
Db	2580 CCTCTATGAGCTGATGCCGGGGAGGCCGAGGCCGC 2633
RESULT	2
AL39517	AL39517 1010 bp mRNA linear EST 16-FEB-2001
LOCUS	AL39517
DEFINITION	prime_mRNA sequence.
ACCESSION	AL39517
VERSION	AL39517.1
KEYWORDS	EST
SOURCE	human.
ORGANISM	Homo sapiens
KARYOTYPE	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 1010)
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
QY	2194 ACCTGTTGAGGTGACAGCCCTCAAGATCTCGTCAGTTCCTTAATGTC 2253
Db	1818 ACCTGTTGAGTAGGAGGCCGCCAAGATCTCGTCAGTTCCTTAATGTC 1877
QY	2254 GTAGGGCACCTTGGTGTCAAGATCTACGGCCAGATGCCACAAATG 2313
Db	1878 ACAAGGACCCCTTGTCTAGTGAAGATCTCCGCAGATGCCACAAATG 1937
QY	2314 CCAGCTTCCTGTCAGTGAAGAGGTGAGATCTGTCAGG 2373
Db	1938 -----CCAGTGTGTTCTGAGGAGTCACTGAGTACATCAGTCACGG 1979
QY	2374 TCAAGGACCCAACTCATGGCTGCTGGGGTGTGAGGACCCCTGTCA 2433
Db	1980 TGAAGGACCCAACTCATGGCTGCTGGGGTGTGAGGACCCCTGTCA 2039
QY	2434 TGATTAAGTACTACATGGAGACGCCGACCTCAACCAAGTCTCACTG 2493
Db	2040 TGATCACTACATGGAGATGGCATGACCAAGTCTCACTG 2099
QY	2494 AGGACAGGAGGCCAGGGGCCCTGGGAGCGGCCAGGTGCGGGGCCACATCA 2553
Db	2100 AGAACAGGCACTCAGGGCTCTGGACACAGAGTGTGACCAAGG 2159
QY	2554 GCTACCCATCTGGTGTGAGCTGGGGCCAGATGCCCTGGCATGGCATGCC 2613
Db	2160 GCTACCCATCTGGTGTGAGCTGGGGCCAGATGCCCTGGCATGCC 2219
QY	2614 CACTCACTTGTACATCGGACCTCGCACCGGAACATCTTGGTTGGGAAATTCA 2673
Db	2220 CGCTGAACTTGTGCAATCGGACCTCGCACCGGAACATCTTGGTTGGGAAATTCA 2279
QY	2674 CCATCAAATTCGCAACTTGGCATGAGCGGAACTCTATGCTGGGGACTTACCGTG 2733
Db	2280 CCATCAAATTCGCAACTTGGCATGAGCGGAACTCTATGCTGGGGACTTACCGTG 2339
Qy	2734 TCGAGGGCGGCACTGCTGCCATCGTAGTGGAGCTGGAGTCATCTCATGGGA 2793
Db	2340 TCCAGGGCGGGCTGCCATCGTAGTGGAGCTGGAGTCATCTCATGGGA 2399
Qy	2794 AGTCACGCTGGCACTGGCTGGGGCTGGGGAGTGTGATG 2853
Db	2400 AGTCACACAGCCAGTGAAGTGGCCCTGGGAGTGTGATG 2459
Qy	2854 TCTTAGGCCAGCCTTGGGAGCTACCGACAGGGTCATCGAGAACGGGGG 2913
Db	2460 TCTCAGGCCAGCCTTGGGAGCTACAGATGAGGTATCGAGAACGGGG 2519
Qy	2914 AGTCCTGGACAGGCCGGAGCTGGGGAGTGTGACCTGGGGCTGGGGAGTGTG 2973
Db	2520 AGTCTCAGGCCAGGCCGGAGTGTGACCTGGGGCTGGGGAGTGTG 2579
Qy	2974 GCCTATGAGCTGATCTGGGGCTGAGCCGGAGCTGACGAGCCAC 3027
Db	2580 CCTCTATGAGCTGATGCCGGGGAGGCCGAGGCCGC 2633
RESULT	2
AL39517	AL39517 1010 bp mRNA linear EST 16-FEB-2001
LOCUS	AL39517
DEFINITION	prime_mRNA sequence.
ACCESSION	AL39517
VERSION	AL39517.1
KEYWORDS	EST
SOURCE	human.
ORGANISM	Homo sapiens
KARYOTYPE	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 1010)
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
QY	2860 GGGCCAGCCCTTGGGAGCTCACGACGAGGTGATCGAGAGACGGGGGGAGTCT 2919
Db	361 GGGCCAGCCCTTGGGAGCTCACGACGAGGTGATCGAGAGACGGGGGGAGTCT 420
QY	2920 TCCGGGACCAAGGCCGGAGGTGTAACCTGTCGCCGCCGCTGCCGCCAGG---CTAT 2979
Db	421 TCCGGGACCAAGGCCGGAGGTGTAACCTGTCGCCGCCGCTGCCGCCAGG---CTAT 478
Qy	2980 ATGAGCTGAGCTGGTGTGAGCCGGAGTCTGAGGCCACCCCTTCCAGC 3039
Db	479 ATGAGCTGAGCTGGTGTGAGGCCACCCCTTCCAGC 538
QY	3040 TGCATGGTCTGGCAGAGATGACTCACACGGTGAATCACACAGCTGGCC 3099
Db	539 TGCATGGTCTGGCAGAGATGACTCACACGGTGAATCACACAGCTGGCC 538
QY	3100 CTCCCTCAGGAGTATCAGGGGAGGCCAGTGGAGTGTGAGCTGATCTG 3159



LCUS AL528664 LTL\_NFL001\_NBC4 987 bp mRNA linear EST 13-FEB-2001  
 DEFINITION mRNA sequence.  
 ACCESSION AL528664  
 VERSION AL528664.1 GI:1292157  
 KEYWORDS ESR  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 987)  
 AUTHORS Li, W.B., Gruber, C., Jerssee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Séquençage  
 BP 191 91066 EVRY cedex - France  
 Email: [sequen@genoscope.cns.fr](mailto:sequen@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).  
 FEATURES source  
 /location/Qualifiers  
 1. 987  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS00001YK21"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /clone="LTL\_NFL001\_NBC4"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RI sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed  
 by Life Technologies. Contact: Feng Liang Life  
 Technologies, a division of Invitrogen 9800 Medical Center  
 Drive Rockville, Maryland 20850, USA Fax : (1) 301 610  
 8371 Email : [fling@lifetech.com](mailto:fling@lifetech.com) URL :  
 http://fulllength.invitrogen.com/ 6 others  
 BASE COUNT 186 a 263 c 332 g 200 t  
 ORIGIN

Query Match 20.6%; Score 814.6; DB 9; Length 987;  
 Best Local Similarity 98.7%; Pred. No. 2.3e-137; Mismatches 3; Indels 4; Gaps 3;  
 Matches 848; Conservative 4; Mismatches 3; Indels 4; Gaps 3;

QY 277 AGGAGATGCTGCCACCCCTTAGGCCGAGGATCATGGAGCTTGGACAGGCC 336  
 Db 90 AGGAGATGCTGCCACCCCTTAGGCCGAGGATCATGGAGCTTGGACAGGCC 149  
 QY 337 TGTATCCTTACTGCTCTGCTTGGGCCAGTGGAATGAGATGGTACATGAGGACATT 396  
 Db 150 TGTATCCTTACTGCTCTGCTTGGGCCAGTGAGATGGTACATGAGGACATT 209  
 QY 397 TGTATCCTGCCAGTGCCCTTGCCCTGGGAGCAGGACGACATCCGAGCTG 456  
 Db 210 TGTATCCTGCCAGTGCCCTGGGAGCAGGACGACATCCGAGCTG 269  
 QY 457 ACACCTCTCCAGCTCTGAGATCCACTGGCCGGCCACAGCAGTTGGAGA 516  
 Db 270 ACATCTCTGCCAGCTCCAGTCCAGTCCACTGGCCGGCCACAGCAGTTGGAGA 329  
 QY 517 GCAGTGCAGGGGATGGGCCCTGGCCGGAGGGAGGTTCAGGGAGGAGT 576  
 Db 330 GCAGTGCAGGGGATGGGCCCTGGCCGGCA-ATGGGTGTTCCCAAGGAGGAGT 387  
 QY 577 ACTGGAGTGCATCACAGGACTCACCTGGCTGGGAGGACCAAGGGAGGC 636  
 Db 388 ACTTCAGTGTGATCTACAGGACTCACCTGGCTGGGAGGACCAAGGGAGGC 447  
 QY 637 ATGCCGGGGCTGGCAAGGAGTCTCCGGAGCTACCGAGTGGTTACTCCGGATG 696  
 Db 448 ATGCCGGGGCTGGCAAGGAGTCTCCGGAGCTACCGAGTGGTTACTCCGGATG 507

RESULT 5  
 BM043544  
 LOCUS BM043544 898 bp mRNA linear EST 07-NOV-2001.  
 DEFINITION mRNA sequence.  
 ACCESSION BM043544  
 VERSION BM043544.1 GI:16772811  
 KEYWORDS EST  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 898)  
 AUTHORS NIH-MGC <http://mgc.ncbi.nih.gov/>,  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgbps1@mail.nih.gov](mailto:cgbps1@mail.nih.gov)  
 Tissue Procurement: BCDP/DP  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LILN)  
 Clone distribution by: Incyte Genomics, Inc.  
 found through the I.M.A.G.E. Consortium/LILN at:  
<http://image.liln.gov>  
 Plate: LILC1926 row: 1 column: 18  
 High quality sequence stop: 837.  
 FEATURES source  
 /location/Qualifiers  
 1. 898  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5445617"  
 /clone\_1b="NIH-MGC\_40"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: prostate; Vector: pOB7; Site\_1: XbaI;  
 Site\_2: EcoRI; cDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XbaI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University

DEFINITION	BT257820	770 bp	mRNA	linear	EST	17-JUL-2001
LOCUS	60396918	NIH_MGC_12	Homo sapiens	cDNA clone	IMAGE:5109436	5',
REFERENCE	60396918	NIH_MGC	sequence.			
ACCESSION	BT257820					
VERSION	BT257820.1	GI:14813571				
KEYWORDS	EST					
SOURCE						
ORGANISM	Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE	1 (bases 1 to 770)					
AUTHORS	NIH-MGC	<a href="http://mgc.ncbi.nlm.nih.gov/">http://mgc.ncbi.nlm.nih.gov/</a>				
TITLE	Unpublished (1999)	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Contact: Robert Strausberg, Ph.D.					
COMMENT	Email: cgapbs-r@mail.nih.gov					
TISSUE	Procurement: ATCC					
CDNA Library Preparation: Life Technologies, Inc.						
CDNA Library Arrayed by: Incyte Genomics, Inc.						
DNA Sequencing by: Incyte Genomics, Inc.						
Clone distribution: MGC clone distribution information can be found through the T.M.A.G.E. Consortium/LLNL at:						
http://image.llnl.gov						
plate: LIAM1265 row: c column: 05.						
High quality sequence stop: 765.						
FEATURES						
Source						
Location/Qualifiers						
1. 770						
/organism="Homo sapiens"						
/db_xref="taxon:9606"						
/clone="IMAGE:5109436"						
/clone_id="NIH_MGC_12"						
/tissue_type="cervical carcinoma cell line"						
/lab_host="DH10B"						
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT Technologies."						
Average insert size 1.4 kb. Library prepared by Life						
BASE COUNT	144 a	207 C	262 g	157 t		
ORIGIN						
Query Match	18.1%	Score	715.8	DB	10	Length
Best Local Similarity	97.3%					770
Matches	750	Conservative	0	Mismatches	17	Indels
					4	Gaps
QY	229	GAAGGTGCTTATCCTACTGACCGATGGGACTGAGTGGGACTGAGTGGAGATGAGAATGCCAGAGATGCTGCC	288			
Db	61	CCCCACCCCTTAGGCCGAGGATCAGGACTATGGACCAGAGGCCCTGTCATCTTAC	120			
QY	289	CCCCACCCCTAGGCCGAGGATCAGGACTATGGACCAGAGGCCCTGTCATCTTAC	348			
Db	1	GAAGGTGCTTATCCTACTGACCGATGGGACTGAGTGGGACTGAGTGGAGATGAGAATGCCAGAGATGCTGCC	60			
QY	349	TGCTGCTGCTCTGGTGGAGATGCTGAGATGAGGAGATTTGATCCCTGCCA	408			
Db	121	TGCGCTGCTGCTCTGGTGGAGATGCTGAGATGAGGAGATTTGATCCCTGCCA	180			
QY	409	AGTGGCGCTATGCCCTGGCATCGAGCGGCCATCCAGACAGTGCACATCTCTGCTT	468			
Db	181	AGTGGCGCTATGCCCTGGCATGGAGGCCGACCATCCAGACAGTGCACATCTCTGCTT	240			
QY	469	CCAGCTCTGTCAGATCCACTGCCGCCAGAGCAGGTGGAGAGGAGCTGACGGGG	588			
Db	241	CCAGCTCTGTCAGATCCACTGCCGCCAGAGCAGGTGGAGAGCAGTGCTGCC	300			
QY	529	ATGGGGCTCTGGTGGCCGGAGGGATGGCTGAGGACGGACCATCCAGACAGTGCACATCTCTGCTT	588			
Db	301	ATGGGGCTCTGGTGGCCGGAGGGATGGCTGAGGACGGACCATCCAGACAGTGCACATCTCTGCTT	360			
QY	589	ATCTACAGACTGCACCTGGTGGCTGAGGACGGACCATCCAGGGAGGAGCTGCTGCC	648			
Db	361	ATCTACAGACTGCACCTGGTGGCTGAGGACGGACCATCCAGGGAGGAGCTGCTGCC	420			

QY 649 TGGCAAGGAGTCTCCGGACACTACGGCTCGTTACTCCGGATGGTCCCGCGTGA 708  
 DB 421 TGGCAAGGAGTCTCCGGACACTACGGCTCGTTACTCCGGATGGTCCCGCGTGA 480  
 QY 709 TGGCAAGGAGTCTCCGGACACTACGGCTCGTTACTCCGGATGGTCCCGCGTGA 768  
 DB 481 TGGCAAGGAGTCTCCGGACACTACGGCTCGTTACTCCGGATGGTCCCGCGTGA 540  
 QY 769 TGGTGTGAAGGACCTGGGCCCATGGTGTGCCGACTGGTCCGCTTCACCCGGG 828  
 DB 541 TGGTGTGAAGGACCTGGGCCCATGGTGTGCCGACTGGTCCGCTTCACCCGGG 600  
 QY 829 CTGACCGGGTCATG---AGTGTCTGCTGCGCGTAGCTGAGCTATGGCTGCCCTGGAGGG 885  
 DB 601 CTGACCGGGTCATGAGCAGTCGACTGCGGTAGAGCTCATGGCTGCCCTGGAGGG 660  
 QY 886 ATGGACTCTGCTTACACGCCCGTGGCCAGACAGTATTCAGGGCTGT 945  
 DB 661 ATGGACTCTGCTTACACGCCCGTGGCCAGACAGTATTCAGGGCTGT 719  
 QY 946 ACTCAACGACTCCACCTATGAGGACATAACGGGACTGAGT 996  
 DB 720 ACTCTAGGACTCCACCTATGAGGACATAACGGGACTGAGT 770

RESULT 7  
 LOCUS B1025684 948 bp mRNA linear EST 04-OCT-2001  
 DEFINITION 603072688FL NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5164480 5',  
 mRNA sequence.  
 ACCESSION B1025684  
 VERSION 1  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 948)  
 AUTHORS NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.  
 VERSION B1025684.1 GI:15937234  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgbps-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 DNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The T.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 MCC clone distribution information can be found through the T.M.A.G.E. Consortium/LLNL at:  
<http://Image.llnl.gov>  
 Plate: LLM1408 row: h column: 17  
 High quality sequence stop: 785.

FEATURES  
 source  
 1. .948  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5164480"  
 /clone\_lib="NIH\_MGC\_119"  
 /tissue\_type="medulla"  
 /lab\_host="B101B"  
 /note="Organ: brain; Vector: PCMV-SPORT6; Site:1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source: normal medulla from  
 anonymous male age 27. Library is oligo-dT primed and  
 directionally cloned (Bcory) site is destroyed upon  
 cloning. Average insert size 1.3 kb. Insert size range  
 0-9-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 013. Note:  
 this is a NIH\_MGC Library."  
 BASE COUNT 178 a 272 c 296 g 202 t  
 ORIGIN

Query Match 18.0%; Score 715; DB 10; Length 948;  
 Best Local Similarity 93.0%; Pred. No. 2, 3e-119;  
 Matches 845; Conservative 0; Mismatches 35; Indels 29; Gaps 8;  
 Matches 845; Conservative 0; Mismatches 35; Indels 29; Gaps 8;

QY 2136 CCTCGATCTGACTCCCTTCAGGAGAGCTGGCGAGGCCAGTTGGGAGGTGAC 2195  
 DB 1 CTCGATCTGACTCCCTTCAGGAGAGCTGGCGAGGCCAGTTGGGAGGTGAC 60  
 QY 2196 CGTGTGAGGTGACGCCCTCAAGTCTGCTCAGTCAGTCAGTTCCCCCTAATGCGT 2255  
 DB 61 CTCGTTGAGGTGACGCCCTCAAGTCTGCTCAGTCAGTCAGTTCCCCCTAATGCGT 120  
 QY 2256 AAGGACACCCCTTGCTGTAGCTGCAAGATCTAGGCGAGATGCCAACAGATGCC 2315  
 DB 179 -----CCAGGATGATTTCCTGAAGAGGTTGAGATCTAGTCGAGCTTGCCGAGGT 222  
 QY 2316 AGCTCTCTGTTCTCCAGGATGATTCTGAAAGAGGAGATCTGCGAGGCTC 2375  
 DB 121 AAGGACACCCCTTGCTGTAGCTGCAAGTCTAGGCGAGATGCCAACAGATGCC 178  
 QY 2376 AAGGACCCCAACATCATTCGCTGCTGGGCTGTGTGCGAGGACCCCTCGCTG 2435  
 DB 223 AAGGACCCCAACATCATTCGCTGCTGGGCTGTGTGCGAGGACCCCTCGCTG 282  
 QY 2436 ATTACGACTACATGGAGAAGGGGACCCCTACCCAGTTCTCAGTGCCTCAC 2495  
 DB 283 ATTACGACTACATGGAGAAGGGGACCCCTACCCAGTTCTCAGTGCCTCAC 342  
 QY 2496 GACAAGGCCAGCGAGGGGCCCTGGAGGGAGGCTGGCGAGGCTGGCTGGAG 2555  
 DB 343 GACAAGGCCAGCGAGGGGCCCTGGAGGGAGGCTGGCGAGGCTGGCTGGAG 402  
 QY 2556 TACCCATGCTGCTCATGCTGCACTGCCACATCSCCTCGCGCATCGCTG 2615  
 DB 403 TACCCATGCTGCTCATGCTGCACTGCCACATCSCCTCGCGCATCGCTG 462  
 QY 2616 CTCACTTGTACATGGACCTGGCCACGGAAACCTCTATGGGAAATTTCAGC 2675  
 DB 463 CTCAACTTGTACATGGACCTGGCCACGGAAACCTCTATGGGAAATTTCAGC 522  
 QY 2676 ATCAAATGGAGCTTGTACATGGACCTGGCCACGGAAACCTCTATGGGAA 2735  
 DB 523 ATCAAATGGAGCTTGTACATGGACCTGGCCACGGAAACCTCTATGGGAA 582  
 QY 2736 CAGGGCGGCCAGTCTGCCATCGCTGATGGCTGCTGGAGTCATCTCAGGGAA 2795  
 DB 583 CAGGGCGGCCAGTCTGCCATCGCTGATGGCTGCTGGAGTCATCTCAGGGAA 642  
 QY 2796 TTCAAGACTCGAGTGACCTGGGGCTTGTTGTGA-CCTCTGGAGGTGATGATC 2854  
 DB 643 TTCAAGACTCGAGTGACCTGGGGCTTGTTGTGA-CCTCTGGAGGTGATC 702  
 QY 2855 CTGTAGGGCCAGCCTTGTGG---CAGTCACCGAGCAGCACTGAGACGGGG 2911  
 DB 703 CTGTAAGGCCAGCTTGTGGAGCTAACCCACGAGGCGAGTCATGAGACGGGG 762  
 QY 2912 GGAGSTCTCGGGGCCAGGGCCGCA-GGTGTCACCTGTCGGGCC -GCCGCCCTGCC 2967  
 DB 763 GGAGTCCTCCCGGCCAGGCCGCAAGTCGTCGCCCTGCGCTTCGCTTC 822  
 QY 2968 CGCAGGGCTA-TATGAGGTGATCTCGCTGGAGGGAGTCAGRCGACCA 3026  
 DB 823 CGCAGGCCATTATGAGCTGATCTGGCTGCA-GACGCCAGT-TGAGCGGGACCA 880  
 QY 3027 CCCTTTCC 3035  
 DB 881 CCTTTCCC 889

RESULT 8  
 LOCUS BG912321 799 bp mRNA linear EST 05-JUN-2001

DEFINITION	60280664F1 NCI_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4939239	QY	2863 CCGAGCCCTTGGCCAGCTACCGACGCCAGGTATGGAGAACGGGGAGTCTCC 2922
5', mRNA sequence.			
ACCESSION	BG312321	Db	421 CCGAGCCCTTGGCCAGCTACCGACGCCAGGTATGGAGAACGGGGAGTCTCC 480
VERSTON	BG312321.1	QY	2923 GGGACCCAGGCCAGGTATGGCCAGCTACCGACGCCAGGTATGGAGAACGGGGAGTCTCC 2981
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		481. GGGACCCAGGCCAGGTATGGCCAGCTACCGACGCCAGGTATGGAGAACGGGGAGTCTCC 540
REFERENCE	1 (bases 1 to 799)	QY	2982 GAGCTGATGCTTGGCAGCTACCGACGCCAGGTATGGAGAACGGGGAGTCTCC 3041
AUTHORS	NIH-MGC http://mgc.ncbi.nih.gov/	Db	541 GAGCTGATGCTTGGCAGCTACCGACGCCAGGTATGGAGAACGGGGAGTCTCC 599
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs@mail.nih.gov		
TISSUE	Procurement: David N. Louis, M.D.		3042 CATGGTTCCTGGCAGGGTCACTCACACGGTGTGATCACACATCCAGCTG--- 3096
CDNA	Library Preparation: Life Technologies, Inc.		Db 690 CATCGGTTCTGGCAGGGTCACTCACACGGTGTGATCACACATCCAGCTG 659
LIBRARY	Arrayed by: The I.M.A.G.E. Consortium (LNL)		3097 CCCCTCCCTAGGGAGGTCCAGGGGAGCCAGTGAACATAAACATCCAGCTG 3154
SEQUENCING	by: Incyte Genomics, Inc.		660 CCCTCCCTAGGGAGGTCCAGGGGAGCCAGTGAACATAAACATCCAGCTG 719
BY	NIH-MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov		QY 3155 TGGACCACTCTGCCCTCCCTCCGACAGCCCATCACCTTAATAGAGGAGCTGAGACTG 3214
PLATE	LLAM0875 row: 0 column: 16		Db 720 TGGACCACTCTGCCCTCCCTCCGACAGCCCATCACCTTAATAGAGGCGTGGAGTGA 779
High quality sequence stop: 768.			QY 3215 CAGGTGGCTGG 3226
FEATURES	Location/Qualifiers		Db 780 CAGGTGGCTGG 791
source			
1. .799			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone_lib=NCI_CGAP_Brn67"			
/tissue_type="anaplastic oligodendrogloma with 1p/19q loss"			
/lab_host="BPH10B (T1 phage-resistant)"			
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.3 Kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."			
BASE COUNT	163 a 238 c 251 g 147 t		
ORIGIN			
Query Match	17.3%	RESULT	9
Best Local Similarity	95.7%	DEFINITION	BE744806 BE744806
Matches	758; Conservative 0; Mismatches 25; Indels 9; Gaps 5;	LOCUS	60157025F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3836046 5', mRNA sequence.
QY	2443 ACTACATGGAGAGGGCAGCTAACAGTTCAGTGAGACAGG 2502	ACCESSION	BE744806 BE744806
Db	1 ACTACATGGAGAGGGCAGCTAACAGTTCAGTGAGACAGG 60	JOURNAL	BE744806.1 GI:10158798
QY	2503 CAGCGAGGGGCCCTGGAGGGCAGCTAACAGTTCAGTGAGACAGG 2562	COMMENT	EST.
Db	61 CAGCCGAGGGGCCCTGGAGGGCAGCTAACAGTTCAGTGAGACAGG 120	KEYWORDS	human.
QY	2563 TCGCGCTCATGGCAGCCACATCGCTCGGCAGCGCTAACAGTCACT 2622	ORGANISM	Homo sapiens
Db	121 TCGCGCTCATGGCAGCCACATCGCTCGGCATCGCTAACAGTCACT 180	REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
QY	2623 TTGTCATCGGGCATGGCAGCGGAACTGCTAGTGGGAATTTCACCATCAA 2682	AUTHORS	1 (bases 1 to 715)
Db	181 TTGTCATCGGGCATGGCAGCGGAACTGCTAGTGGGAATTTCACCATCAA 240	JOURNAL	NIH-MGC http://mgc.ncbi.nih.gov/
QY	2683 TCGCAGACTTGGCATGGCAGCGGAACTGCTAGTGGGAATTTCACCATCAA 2742	COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D.
Db	241 TCGCAGACTTGGCATGGCAGCGGAACTTACCGTGGCTGGAGGCC 300	TISSUE	Procurement: Ling Hong/Rubin Laboratory
QY	2743 GGCGAGTGTGGCATCGCTGGATGGCTGGAGTCATCCATGGGAATTTCACCATCAA 2802	CDNA	Library Preparation: Ling Hong/Rubin Laboratory
Db	301 GGCGAGTGTGGCATCGCTGGATGGCTGGAGTCATCCATGGGAATTTCACCATCAA 360	LIBRARY	Arrayed by: The I.M.A.G.E. Consortium (LNL)
QY	2803 CTGGAGTGTGGCTGGCCTTGGTGTGACCTGGGGAGGTGGCTGATGCTCTGTAGG 2862	DNA	Sequencing by: Incyte Genomics, Inc.
Db	361 CTGGAGTGTGGCTGGCCTTGGTGTGACCTGGGGAGGTGGCTGATGCTCTGTAGG 420	SEQUENCING	NIH-MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
High quality sequence stop: 713.		PLATE	Plate: LLM519 row: a column: 07
FEATURES	Location/Qualifiers		High quality sequence stop: 713.
source			
1. .715			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:3036046"			
/clone_lib="NIH_MGC_9"			
/tissue_type="adenocarcinoma cell line"			
/lab_host="BPH10B (phage-resistant)"			
/note="Organ: ovary; Vector: pOR7; Site_1: XbaI; Site_2: EcoRI; cDNA made by oligo-dT priming. cloned into EcoRI/XbaI sites using the following 5' adaptor: GGACGAG (G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit and Stratagene) and superscript II RT (Life Technologies)."			
BASE COUNT	130 a 226 c 212 g 147 t		
ORIGIN			



QY 3057 GAGGATGCACTCACA-CGGTGTGATCACACATCCAGCTGCCCTCCCTCAGGGAGTGA 3115  
 Db 659 GAGGATGCACTCACAACCCGGTGTGATCACACATCCAG-TGGCTCCCTCAGGGAGTGA 717

QY 3116 TC 3117  
 Db 718 TC 719

RESULT 11

LOCUS BT193181 744 bp mRNA linear EST 10-JUL-2001  
 DEFINITION mRNA sequence.  
 ACCESSION BT193181  
 VERSION BT193181.1 GI:14648201  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 744)  
 AUTHORS NIH-MGC http://mgc.ncbi.nih.gov/  
 TITLE Unpublished (1999)  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.llnl.gov>  
 Plate: L1C14852 row: b. column: 16  
 High quality sequence stop: 675.

FEATURES source

1. .744 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5030223"  
 /clone\_id="NIH\_MGC\_42"  
 /tissue\_type="epithelial carcinoma cell line"  
 /lab\_host="B110B (phage-resistant)"  
 /note="Organ: pancreas; Vector: pOTB7; Site\_1: XbaI;  
 Site\_2: EcoRI; CDNA made by oligo-dT priming.  
 Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGCACGAGTCG. Size-selected >500bp following 5' adaptor: GGCACGAGTCG. Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH-MGC Library."|

BASE COUNT 149 a 212 c 229 g 154 t

ORIGIN

Query Match 17.0%; Score 672; DB 10; Length 744;  
 Best Local Similarity 97.5%; Pred. No. 1.4e-11; Mismatches 7.14; Conservative 0; Mismatches 15; Indels 3; Gaps 3;

REFERENCE 1 (bases 1 to 718)  
 AUTHORS NIH-MGC http://mgc.ncbi.nih.gov/  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Tissue Procurement: James Clever, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.llnl.gov>  
 Plate: L1A10696 row: j column: 22  
 High quality sequence stop: 717.

FEATURES source

1. .718 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

QY 2417 GGACGACCCCTCTGCTATGTTACTGACTACATGGAGAAGGGGACCTAACCGATCTC 2476  
 Db 182 GGAGGACCCCTCTGCTATGTTACTGACTACATGGAGAAGGGGACCTAACCGATCTC 241

QY 2477 CAGTGCCTCCACGTGGAGGACAGGGACCCGGAGGGGCCCTGGGGAGGGCAGGGC 2536  
 Db 242 CAGTGCCTCCACGTGGAGGACAGGGACCCGGAGGGGCCCTGGGGAGGGCAGGGC 301

QY 2537 GCAGGGCCACCATCAGGTACCAATGGTCCTGAGTGGCAGCCAGATGCTCCGG 2596  
 Db 302 GCAGGGCCACCATCAGGTACCAATGGTCCTGAGTGGCAGCCAGATGCTCCGG 361

QY 2597 CATGGCTATCTGCCACACTCAACTTGTACATGGGCCACCGGAACATGCTC 2656  
 Db 362 CATGGCTATCTGCCACACTCAACTTGTACATGGGCCACCGGAACATGCTC 421

QY 2657 AGTGGGAAATTCACTCAAATGCCAGTTGCATAGGCCGAACTCTAGC 2716  
 Db 422 AGTGGGAAATTCACTCAAATGCCAGTTGCATAGGCCGAACTCTAGC 481

QY 2717 TGGGACTTTACCGTCGAGGGGGCAGTGACGCTGCCCACCGGAACTCTAGC 2776  
 Db 482 TGGGACTTTACCGTCGAGGGGGCAGTGACGCTGCCCACCGGAACTCTAGC 540

QY 2777 GTGGACTCTCATGGGAAGTCACTAAATGCCAGTTGCATAGGCCGAACTCTAGC 2836  
 Db 541 GTGCATCTCATGGGAAGTCACTAAATGCCAGTTGCATAGGCCGAACTCTAGC 600

QY 2837 GTGGGAGGCTGCTGCTGAGGGCCAGCCTTGGCACCTCACGAGCAGT 2896  
 Db 601 GTGGGAGGCTGCTGCTGAGGGCCAGCAGTGGCTGAGGACAG-CCTTGGCACAGTCACCGACGAGCT 659

QY 2897 CATTGAGAAG-CGGGGAGTTCGGACAGGGGGAGTGTGACCTGCCCAGC 2955  
 Db 660 CATTGAGAAGCAGGGGGAGTTCGGGAGAGGGGGAGGGGGAGGTGACCTGCCCAGC 719

QY 2956 CGCTGCCTGCC 2967  
 Db 720.CGCCTGCCTGCC 731

RESULT 12

LOCUS BG696171 718 bp mRNA linear EST 07-MAY-2001  
 DEFINITION 602692851 NCI\_CGAP\_Skn3 Homo sapiens cDNA clone IMAGE:4802421 5', mRNA sequence.  
 ACCESSION BG696171  
 VERSION BG696171.1 GI:13961044  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 718)  
 AUTHORS NIH-MGC http://mgc.ncbi.nih.gov/  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Tissue Procurement: James Clever, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.llnl.gov>  
 Plate: L1A10696 row: j column: 22  
 High quality sequence stop: 717.

FEATURES source

1. .718 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

ORGANISM									
Homo sapiens									
/clone="IMAGE:4802421"									
/lab_host="DH10B (T1 phage-resistant)"									
/note="Organ: skin; Vector: pCMV-SPORT6; site_1: NotI; site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size: 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP library."									
BASE COUNT	133 a	228 c	210 g	147 t	ORIGIN				
Query Match	16.7%	Score 662.6;	DB 10;	Length 718;					
Best Local Similarity	98.7%	Pred. No. 6.9e-110;							
Matches	710;	Conservative	0;	Mismatches	4;	Indels	5;	Gaps	4;
QY	2670	TTCACCATCAAATGCGACACTTGGCATAGGCCGAACTCTATGCTGGGACATTAC	2729						
Db	1	TTCACCATCAAATGCGACACTTGGCATAGGCCGAACTCTATGCTGGGACATTAC	60						
QY	2730	CGTGTGCAAGGCCGCACTGCTGCCATCGCTGATGCCCTGGAGTCATCTCAG	2789						
Db	61	CGTGTGCAAGGCCGCACTGCTGCCATCGCTGATGCCCTGGAGTCATCTCAG	119						
QY	2790	GGGAAGTGTACGACTGCGGAGTGAGCTGGCTGGCCCTGGGTGACCCCTGGGAGCTTAC	2849						
Db	120	GGGAAGTGTACGACTGCGGAGTGAGCTGGCTGGCCCTGGGTGACCCCTGGGAGCTTAC	179						
QY	2850	ATGCCTGTAGGGCCAGGCCCTGGCACTCACCAGAGCAGACCG	2909						
Db	180	ATGCCTGTAGGGCCAGGCCCTGGCACTCACCAGAGCAGACCG	239						
QY	2910	GGGAGTGTTCGGGACAGGGCGCGCTGGCTACCTCTCCGGCGCTGCCCTGCCG	2969						
Db	240	GGGAGTGTTCGGGACAGGGCGCGCTGGCTACCTCTCCGGCGCTGCCG	299						
QY	2970	CAGGCCCTATAGCTGCTGGCTGGCTGGGAGCTGAGAGGACCC	3029						
Db	300	CAGGCCCTATAGCTGCTGGCTGGGAGCTGAGAGGACCC	359						
QY	3030	TTTGTCCACGTGCTGGCTGGGAGCTGAGAGGACCC	3089						
Db	360	TTTGTCCACGTGCTGGCTGGGAGCTGAGAGGACCC	419						
QY	3090	CCAGCTGCCCTGGCTGGGAGCTGAGAGGACCC	3149						
Db	420	CCAGCTGCCCTGGCTGGGAGCTGAGAGGACCC	479						
QY	3150	CACATGCCACTCTGCCCTCCGACAGCCATCACCTCTAAAGAGGA	3209						
Db	480	CACATGCCACTCTGCCCTCCGACAGCCATCACCTCTAAAGAGGA	539						
QY	3210	GACTGCAAGGGGGCTGGCCACCCAGGGAGCTGAGCCCT	3267						
Db	540	GACTGCAAGGGGGCTGGCCACCCAGGGAGCTGAGCCCT	599						
QY	3268	ACACTCTCATGTCGCCCTGCTCTCTCTCTAGAGCCCTGCGCCACCCAGCT	3326						
Db	600	ACACTCTCATGTCGCCCTAATCTCTCTCTAGAGCCCTGCGCCACCCAGCT	659						
QY	3327	GGCTCTGTGGAGGGATCTCTCACCTCTCTAGAGCCCTGCGCCACCCAGCT	3384						
Db	660	GGCTCTGTGGAGGGATCTCTCACCTCTAGAGCCCTGCGCCACCCAGCT	718						
RESULT	13								
AL043251	AL043251	772 bp mRNA linear EST 29-FEB-2000							
LOCUS	DKFZP434K0223	r1 434 (synonym: htes3) Homo sapiens cDNA clone							
DEFINITION	DKFZP434K0223_5'	mRNA sequence.							
ACCESSION	AL043251								
VERSION	AL043251.1	GI:5422649							
KEYWORDS	EST, human.								
SOURCE									
MPIS									
Am Klopferspitze 18a D-82152 Martinsried, Germany									
This is the 5' sequence of the clone insert.									
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;									
sequenced by LMD (Ludwig Maximilians University Munich/Germany) within the CDNA sequencing consortium of the German Genome Project.									
No 51 sequence available.									
This clone (DKFZP434K0223) is available at the RZPD in Berlin.									
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.									
FEATURES									
source									
1.772									
/organism="Homo sapiens"									
/db_xref="taxon:9506"									
/clone="DKFZP434K0223"									
/tissue_type="testis"									
/dev_stage="adult"									
/lab_host="DH10B"									
/note="vector: pSPORT1; site_1: NotI; site_2: SalI"									
BASE COUNT	118 a	250 c	260 g	143 t	1 others				
ORIGIN									
Query Match	16.7%	Score 660.2;	DB 9;	Length 772;					
Best Local Similarity	92.4%	Pred. No. 1.9e-109;							
Matches	750;	Conservative	0;	Mismatches	4;	Indels	58;	Gaps	3;
QY	64	CTCTGGAGCCGCCCTCCGACACCGAGCCGGCCGCTCCGGCCCGCCGCTCCG	123						
Db	1	CTCTGGAGCCGCCCTCCGACACCGAGCCGGCCGCTCCGGCCCGCCGCTCCG	60						
QY	124	GCTCTGGCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	183						
Db	61	GCTCTGGCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	120						
QY	184	GGTGTGGACCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	243						
Db	121	GGTGTGGACCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	157						
Db	158	-----AGAGATGCTGCCCTCACCCCTAGAGCTGAGCTGAGCTGAGCTGAG	184						
QY	244	CTGACCGATGGGGTGGACTGAGGAAGGCCAGAGATGCTGCCCTACCCCTAGGC	303						
Db	185	CTGACCGATGGGGTGGACTGAGGAAGGCCAGAGATGCTGCCCTACCCCTAGGC	244						
QY	304	CCGAGGGATCAGGACCTATGGGACAGGAGCCCTTCAGCTGGCTGGCTGGCT	363						
Db	305	CCGAGGGATCAGGACCTATGGGACAGGAGCCCTTCAGCTGGCTGGCTGGCT	364						
QY	484	ATTCACAGCCGCCGACAGCAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	543						
Db	365	ATTCACAGCCGCCGACAGCAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	424						
QY	544	CCGAGGGGGCTTCCAAAGGGAGGAGTCTGGCTGGCTGGCTGGCTGGCTGGCT	603						
Db	425	CCGAGGGGGCTTCCAAAGGGAGGAGTCTGGCTGGCTGGCTGGCTGGCTGGCT	484						

QY 604 ACTGGTGGCTCTGGGGACCCAGGAGGGATGCCAGGGCTGGCAAGGAGTCT 653  
 Db 485 ACTGGTGGCTCTGGGGACCCAGGAGGGATGCCAGGGCTGGCAAGGAGTCT 544  
 QY 664 CCCGAGTACCGGCTGCGTACATCCCGGATGATGCGCTGAGGAC 723  
 Db 545 CCCGAGTACCGGCTGCGTACATCCCGGATGATGCGCTGAGGAC 604  
 QY 724 GCTGGTCAAGGAGTGATCTCAGCANTGAGAACCTGAGGAGTGGGCTGAGGAC 783  
 Db 605 GCTGGTCAAGGAGTGATCTCAGCANTGAGAACCTGAGGAGTGGGCTGAGGAC 654  
 QY 784 TTGGCCGCCAAGTGGCCGACTGGTGCCTACCCCCGGCTGACGGCTCATGA 843  
 Db 665 TTGG-CCCCCATGCTGGCGAGTGGCGCTCTACCCCCGGCTGACGGCTCATGA 723  
 QY 844 GTGTCGTCGCGGGTAGAGCTCATGGCTGC 875  
 Db 724 GCGCTGTCGCGGGTAGA-CTCTATGGCTGC 754

RESULT 14  
 BE304984/c  
 LOCUS BE304984 912 bp mRNA linear EST 13-JUL-2000  
 DEFINITION mRNA sequence.  
 REFERENCE BE304984.1  
 ACCESSION BE304984  
 VERSION EST  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammal; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 912)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgsbs-r@mail.nih.gov  
 Tissue Procurement: NCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The T.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: NIH Intramural Sequencing Center  
 Clone distribution: MGC clone distribution information can be  
 found through the T.M.A.G.E. Consortium/LNL at: image.lnl.gov  
 Plate: LNCM51 row: k column: 08  
 High quality sequence start: 9  
 High quality sequence stop: 702.  
 Location/Qualifiers  
 FEATURES source  
 1. .912  
 /organism="Homo sapiens"  
 /clone="IMSR:299615"  
 /clone\_id="NIH\_MGC\_15"  
 /clone\_type="adenocarcinoma cell line"  
 /lab\_host="DHLB (phage-resistant)"  
 /note="Organ: colon; Vector: pORN7; Site\_1: XbaI; Site\_2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XbaI sites using the following 5'  
 adaptor: GCGAGGAG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and superscript II RT (Life Technologies)"  
 BASE COUNT  
 ORIGIN 211 a 244 c 277 g 180 t  
 Query Match 16.5%; Score 655.6; DB 10; length 912;  
 Best Local Similarity 95.4%; Pred. No. 1.3e-108; Matches 740; Conservative 0; Mismatches 29; Indels 7; Gaps 6;  
 REFERENCE 1  
 AUTHORS Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsuura,A., Bala,G.S., Simpson,D.H.,  
 Brunsden,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,  
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

QY 3098 CCCTCCCTAGGGACTGATCCAGGGAGACAC--TAAACAGAGGACAT 3155  
 Db 778 CCCTCCCTAGGGTCGATCCAGGAAGCAGCATGCTAACTAGGATCACAT 719  
 QY 3156 GGCACCTCTCCCTTCCCTCCCGACAGGACATCACCTTAATAGGGAGTGC 3215  
 Db 718 CGCACCTCGCCGTCCTCCCGACAGGACATCACCTTAATAGGGAGTGC 659  
 QY 3216 AGGGGGCGGGCCACCCAGGGACTGAGGCGCCCTTCCTGACACACTTC 3275  
 Db 658 AGGAGGGCTGGGCCACCCAGGTACTGA-GCCCTTCGCGCCCTGCGACACTTC 600  
 QY 3334 TGGATGGGATCTCTCACCCCTCTACCCATCCCTGGGAAGGGTGGGAGAAT 3393  
 Db 539 TGGATGGGATCTCTCACCCCTCTACCCATCCCTGGGAAGGGTGGGAGAAT 480  
 QY 3394 AGGATAGACTGGACATGC-CCATGGAGCACCTGGGCCACTGGACACACTGAT 3452  
 Db 479 AGGATAGACTGGACATGC-CCATGGAGCACCTGGGCCACTGGACACACTGAT 420  
 QY 3453 CCTGGAGAGGGCTGGCTCG--CCCCACCTTCTCCCTCACACTGACCCACTGG 3511  
 Db 419 CCTGGAGAGGGCTGGCTCG--CCCCACCTTCTCCCTCACACTGACCCACTGG 360  
 QY 3512 CTGAGATCTGGGGTGAAGGAGAACAGAGGAGAAATGTTCTCTGGCTGTC 3571  
 Db 359 CTGAGAATCTGGGGTGAAGGAGAACAGAGGAGAAATGTTCTCTGGCTGTC 300  
 QY 3572 CTGACTCTCTCTAGCTGGCTCTCTCTCCATACCTGAAACACTGGACCTGG 3631  
 Db 299 CTGACTCTCTCTAGCTGGCTCTCTCCATACCTGAAACACTGGACCTGG 240  
 QY 3632 GGTACCCGCCCGCCGCGCTCAGTCACCCCACTTCACCTGGAGCTGTGACTGAG 3691  
 Db 239 GGTACCCGCCCGCCGCGCTCAGTCACCCCACTTCACCTGGAGCTGTGACTGAG 180  
 QY 3692 CTCTCTAACCTATAGCTTCTGGAGTAATTTGGATGGGGAAAGGGAGC 3751  
 Db 179 CTCTCTAACCTATAGCTTCTGGAGTAATTTGGATGGGGAAAGGGAGC 120  
 QY 3752 AACGCCCTCATGCCCTGGGCTGGGAGATCTCTAGTGACTGCTGCACATGATTTCTAT 3811  
 Db 119 AACGCCCTCATGCCCTGGGCTGGGAGATCTCTAGTGACTGCTGCACATGATTTCTAT 60  
 QY 3812 AATCTCTGGGTTGATTTGGGGAGAGACAGATTTTACACTAATA 3867  
 Db 59 AATCTCTGGGTTGATTTGGGGAGAGACAGATTTTACATAATA 4

RESULT 15  
 BF944646 BF944616 662 bp mRNA linear EST 22-JAN-2001  
 LOCUS BF0-NNT171-181000-001-c05 NN171 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION NN171 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF944646  
 VERSION BF944646.1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammal; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 662)  
 AUTHORS Diaz Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,  
 Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsuura,A., Bala,G.S., Simpson,D.H.,  
 Brunsden,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,  
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.

JOURNAL COMMENT	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
Contact: Simpson A.J.G.	
Laboratory of Cancer Genetics	
Ludwig Institute for Cancer Research	
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil	
Tel: +55-11-2707001	
Fax: +55-11-2704922	
Email: asimpson@ludwig.org.br	
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL ( <a href="http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PM05t2=PM0-NN1171-181000-001-c05&amp;t3=2000-10-18&amp;t4=1">http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PM05t2=PM0-NN1171-181000-001-c05&amp;t3=2000-10-18&amp;t4=1</a> )	
Seq primer: puc 18 forward	
High quality sequence/Stop: 660.	
FEATURES	Location/Qualifiers
source	1. - 662 /organism="Homo sapiens" /db_xref="txon:9606" /clone lib="NM1171" /dev_stage="Adult" /note="Organ: nervous-normal; Vector: puc18; Site 1: Smal; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions." BASE COUNT ORIGIN
	124 a 214 c 191 g 132 t 1 others
Query Match	16.5%; Score 654.6; DB 10; Length 662;
Best Local Similarity	99.2%; Pred. No. 1.9e-108;
Matches	657; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy	2798 CACGACTGGGAGAGTGAAGTGTGGGGCTTGTGGTGAACCCCTGGGGAGCTGCTCTG 2857
Db	1 CACGACTGGGAGAGTGAAGTGTGGGGCTTGTGGTGAACCCCTGGGGAGCTGCTCTG 60
Qy	2858 TGGGGCCAGCCATTGGGAGGTACCCACGGCAGGCTCATCGAAACCGGGGAGTT 2917
Db	61 TAGGGCCAGCCATTGGGAGGTACCCACGGCAGGCTCATCGAAACCGGGGAGTT 120
Qy	2918 CTTCGGGACCAAGCCGGCAGGTACCTGTCCGGCCGCTGCTGGCCGGGGCT 2977
Db	121 CTTCGGGACCAAGCCGGCAGGTACCTGTCCGGCCGCTGCTGGCCGGGGCT 180
Qy	2978 ATATGAGGTGATGCTTCGGCTGGAGCGGGAGCTGAGCAAGGACCCCTTTCGA 3037
Db	181 ATATGAGGTGATGCTTCGGCTGGAGCGGGAGCTGAGCAAGGACCCCTTTCGA 240
Qy	3038 GCTGCATGGTCTGGAGAGATGCACTAACACGGGTGATCACATCAGCTGC 3097
Db	241 GCTGCATGGTCTGGAGAGATGCACTAACACGGGTGATCACATCAGCTGC 300
Qy	3098 CCTTCCTCAGGGAGTGTCAAGGGAGCCAGTGAACATAAACAAAGGAGAACATGG 3157
Db	301 CCTTCCTCAGGGAGTGTCAAGGGAGCCAGTGAACATAAACAAAGGAGAACATGG 350
Qy	3158 CACCTCTGGCTTCCCTCCGAGAGCCATCCTTAATAGGGCACTGAGCTGCAG 3217
Db	361 CACTCTGGCTTCCCTCCGAGAGCCATCCTTAATAGGGCACTGAGCTGCAG 420
Qy	3218 GTCAGCTGGGCAACCCAGGAGCTGAGCCCTTCTGGCACACTCTCAT 3277
Db	421 GTCAGCTGGGCAACCCAGGAGCTGAGCCCTTCTGGCACACTCTCAT 480
Qy	3278 GTCAGCTGGGCAACCCAGGAGCTGAGCCCTTCTGGCACACTCTCAT 3337
Db	481 GTCAGCTGGGCAACCCAGGAGCTGAGCCCTTCTGGCACACTCTCAT 540

Qy	3338 TGGATCTCCACCCCTCTAGCCATCCCTGGAGAGGTGGGAGAAATAGGA 3397
Db	541 TGGATCTCCACCCCTCTAGCCATCCCTGGAGAGGTGGGAGAAATAGGA 600
Qy	3398 TAGACATGGACATGGCCATGGAGACACAGTGAATCTGG 3457
Db	601 TAGACATGGACATGGCCATGGAGACACAGTGAATCTGG 660
Qy	3458 AG 3459
Db	661 AG 662

Search completed: October 5, 2002, 16:12:29  
Job time: 8690 sec